

РОЗПІЗНАВАННЯ ЗОБРАЖЕНЬ ЗА ДОПОМОГОЮ СІАМСЬКОЇ НЕЙРОННОЇ МЕРЕЖІ

*Петровський А. А., студент; Могильний С. Б., к. т. н., доцент
Національний технічний університет України «Київський політехнічний
інститут імені Ігоря Сікорського», м. Київ, Україна*

Зараз, в наш час, питання безпеки є найбільш важливим в інформаційному суспільстві. Тому необхідно використовувати різні методи захисту. А для доступу до даних необхідно використовувати аутенфікацію — тобто переконатися, що доступ отримає дійсно та особа, яка має доступ.

До раніше створеної системи Інтернету-речей було вирішено додати авторизацію за допомогою обличчя. Для цього було вирішено розгорнути ти навчити нейрону мережу.

Головним в нейронній мережі є питання архітектури. Щодо архітектури нейронної мережі було вирішено поділити розпізнавання на дві частини:

- Визначення власне обличчя на зображенні
- Розпізнавання обличчя

Для визначення обличчя потрібно було зробити такі дії:

1. Скористатися алгоритмом напрямлених градієнтів — HOG (Histogram of Oriented Gradients) для визначення обличчя
2. Власне розпізнавання обличчя

Перший пункт можна здійснити за допомогою бібліотеки dlib, яка містить вже навчену модель для розпізнавання обличчя. За допомогою цієї бібліотеки можна лише виділити обличчя.

Алгоритм HOG є окремою нейронною мережею, яка навчалась на датасеті, який містить близько кількох мільйонів зображень на яких присутні обличчя в різних умовах або повністю відсутні. Сам принцип роботи полягає у конвертації самого зображення у відтінки сірого. Зміною відтінків називається градієнтом. Якщо зміна кольору відбувається з темного до світлішого, то це називається позитивним градієнтом, в іншому випадку — негативним. Зміна з справа на ліво називається горизонтальним градієнтом, а зверху вниз — вертикальним. Далі відбувається зчитування матриці градієнтів з матрицею розміром 8x8. Самі градієнти визначаються, як різниця сусідніх пікселів [1].

Моя розробка полягає у створенні алгоритму розпізнавання обличчя.

Дана нейрона мережа має сіамську архітектуру, тобто навчаються паралельно дві гілки, які потім об'єднуються в одну. Вона містить 37 шарів і 269,741,995 параметрів для процесу навчання. Розроблена вона за допомогою фреймворку для машинного навчання TensorFlow в поєднанні з Keras та мови програмування Python. Навчання проводилося, використовуючи

хмарні технології Kaggle, оскільки звичайні персональні комп'ютери не можуть забезпечити необхідну потужність для навчання такої архітектури мережі.

Сама архітектура моделі виглядає так:

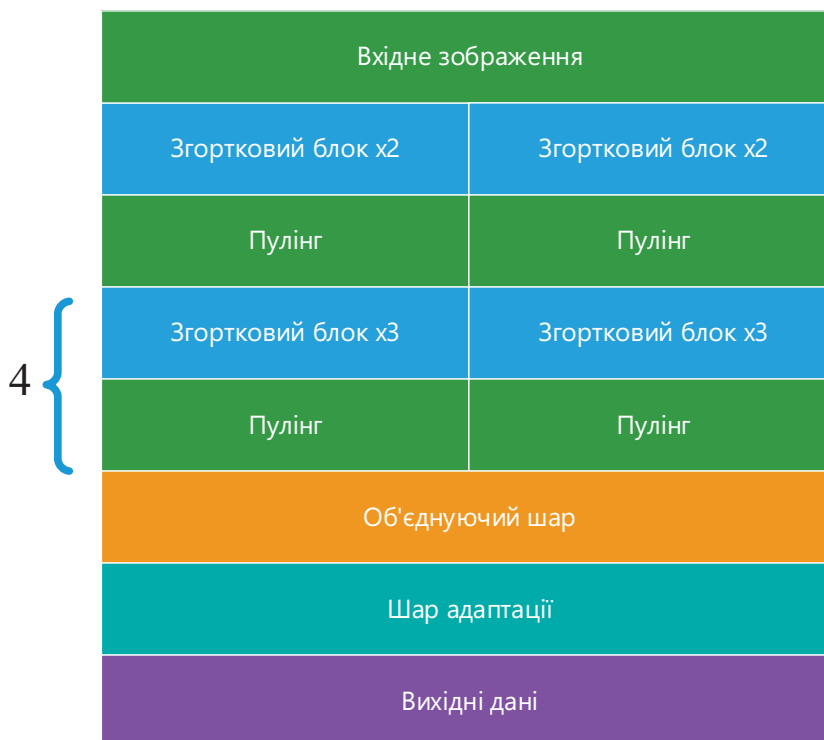


Рисунок 1. Архітектура нейронної мережі

За основу було взяту модель VGG16, яка була раніше розроблена в Масачусетському технологічному університеті. Початкова модель має таку архітектуру: на вхід подається зображення розмірами 224x224. Далі зображення проходить через стек згорткових шарів, розмір вікна яких 3x3. Ядро згортки спочатку вибирається випадковим, а потім підлаштовується за допомогою методу зворотного поширення. Після кожного згорткового шару є зменшення розміру матриці шляхом вибору найбільшого числа з вікна [2].

Далі я змінив цю нейронну мережу, зробивши її сіамською, тобто замість однієї мережі будуть паралельно навчатися 2 нейронні мережі описані вище, які потім об'єднуються в одну. Об'єднання цих двох нейронних мереж можливо за допомогою шару Concatenate. Крім цього було додано ще кілька згорткових шарів для більш детальної обробки параметрів та кілька шарів після об'єднання для забезпечення правильної обробки зображень. Потім йде процес класифікації: для цього додається ще 2 шари з 4096 каналів та 1 шар з n каналів, де n — кількість зображень в датасеті для навчання.

Нижче наведені графіки навчання:

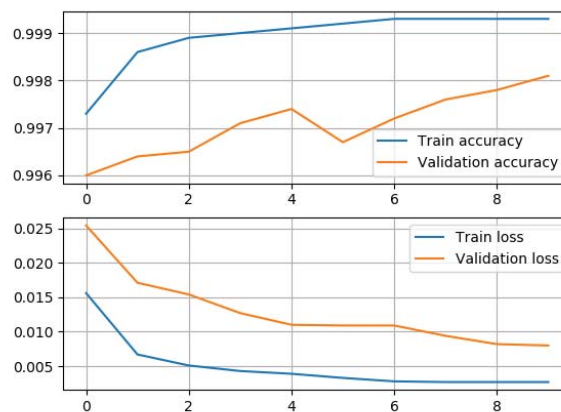


Рисунок 2. Графік ефективності навчання

Як видно з графіку, який наведений нижче точність під час тренування зросла на 0.2% — для машинного навчання це є досить великим показником при наближенні до 100%. Також можна побачити, що внаслідок навчання моєї моделі втрати складають менше 1%.

Перелік посилань

1. P. Felzenszwalb. Efficient matching of pictorial structures / P. Felzenszwalb. // CVPR. – 2000. – С. 66–75.
2. D. A. Forsyth. Probabilistic methods for finding people / D. A. Forsyth, S. Ioffe. // IJCV. – 2001. – С. 45–68.

Анотація

В даній публікації представлена архітектура власної нейронної мережі для розпізнавання обличь. Дана архітектура була розроблена на основі існуючої нейронної мережі, створеної в Масачусетському технологічному університеті, VGG16. Дана нейронна мережа складається з 37 шарів та є сіамською. Точність даної нейронної мережі складає 99.8%, а втрати — менше 1%.

Ключові слова: штучний інтелект, нейронна мережа, розпізнавання обличь.

Аннотация

В данной публикации представлена архитектура нейронной сети для распознавания лиц. Данная архитектура была разработана на основе уже существующей нейронной сети, созданной в Массачусетском технологическом университете, VGG16. Данная нейронная сеть имеет в себе 37 слоев и есть сиамской. Точность этой нейронной сети составляет 99.8%, а потери — меньше 1%.

Ключевые слова: искусственный интеллект, нейронная сеть, распознавание лиц.

Abstract

In this publication architecture of neural network for face recognition is presented. This architecture has been developed using existing neural network, created at Massachusetts institute of technologies, VGG16. This neural network has 37 layers and it is siamese. It has precision about 99.8% and loss — lower than 1%.

Keywords: artificial intelligence, neural network, face recognition.